

# Notice of Allowability

Application No.

10/826,162

Examiner

JEAN B. FLEURANTIN

Applicant(s)

ZHANG ET AL.

Art Unit

2162

## -- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to 04/25/07.
2. ☒ The allowed claim(s) is/are 1-5, 7-17, 19-29, 31-39 and 41-45.
3. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
  - a) ☐ All    b) ☐ Some\*    c) ☐ None    of the:
  1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

\* Certified copies not received: \_\_\_\_\_.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.

**THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.**

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
  5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
    - (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      - 1) ☐ hereto or 2) ☐ to Paper No./Mail Date \_\_\_\_\_.
    - (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date \_\_\_\_\_.
- Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

### Attachment(s)

1. ☒ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3. ☒ Information Disclosure Statements (PTO/SB/08),  
Paper No./Mail Date See Continuation Sheet
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☐ Interview Summary (PTO-413),  
Paper No./Mail Date \_\_\_\_\_
7. ☐ Examiner's Amendment/Comment
8. ☒ Examiner's Statement of Reasons for Allowance
9. ☐ Other \_\_\_\_\_

  
Jean Bolte Fleurantin  
Patent Examiner  
Technology Center 2100

Continuation of Attachment(s) 3. Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date: 11/03/06; 12/19/06; 2/12/07 & 4/25/07.

#### **DETAILED ACTION**

1. This is in response to Applicant's communication on 04/25/07.

The following is the status of claims:

Claims 6, 18, 30 and 40 have been canceled as indicated in the examiner's amendment.

Claims 1-5, 7-17, 19-29, 31-39 and 41-45 remain pending for examination.

#### *Information Disclosure Statement*

The information disclosure statement (IDS) submitted on 11/03/06, 12/19/06, 2/12/07 and 4/25/07. The submission is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

#### *Response to Arguments*

Applicant's arguments, filed 2/12/07, with respect to all pending claims have been fully considered and are persuasive. Therefore, the rejection of last Office action has been withdrawn.

### EXAMINER'S AMENDMENT

An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with Ms. Shirley Anderson (Reg. No. 57,763) on April 25, 2007.

The application has been amended as follow:

In the claims,

Claims 6, 18, 30 and 40 have been cancelled.

Replace old versions of claims 1, 13, 25 and 36 with new versions

Claim 1

A method for verifying relevance between terms and Web site contents, the method comprising:  
retrieving site contents from a bid URL;

formulating expanded term(s) comprising at least one of semantically or contextually related to bid term(s), which are mined from a search engine in view of high-frequency of occurrence historical query terms;

generating content similarity and expanded similarity measurements from respective combinations of the bid term(s), the site contents, and the expanded terms, wherein the similarity measurements indicate relatedness between respective ones of the bid term(s), site contents, or expanded terms;

calculating category similarity measurements between the expanded terms and the site contents in view of a similarity classifier, wherein the similarity classifier has been trained from mined web site content associated with directory data;

calculating a confidence value from combined ones of multiple similarity measurements, wherein the combined ones comprise content, expanded, and category similarity measurements, wherein the confidence value provides objective measure of relevance between the bid term(s) and the site contents;

analyzing the confidence value to identify the bid term(s); and

using the bid term(s) identified to increase traffic to a site to obtain site exposure;

wherein generating the category similarity measurements further comprises:

extracting features from Web site content associated with the directory data, the features comprising a combination of at least one of title, metadata, body, hypertext link(s), visual feature(s), and summarization by page layout analysis information;

reducing dimensionality of the features via feature selection;

categorizing the features via a classifier model to generate the similarity classifier;

generating respective term vectors from the bid term(s), the site contents, and the expanded terms; and

calculating similarity between the respective term vectors as a function of the similarity classifier to determine the category similarity measurements.

#### Claim 13

A computer-readable storage medium comprising computer-executable instructions for verifying relevance between terms and Web site contents, the computer-executable instructions comprising instructions for:

retrieving site contents from a bid URL;

formulating expanded term(s) comprising at least one of semantically or contextually related to bid term(s), which are mined from a search engine in view of high-frequency of occurrence historical query terms;

generating content similarity and expanded similarity measurements from respective combinations of the bid term(s), the site contents, and the expanded terms, wherein the similarity measurements indicate relatedness between respective ones of the bid term(s), site contents, or expanded terms;

calculating category similarity measurements between the expanded terms and the site contents in view of a similarity classifier, wherein the similarity classifier has been trained from mined web site content associated with directory data;

calculating a confidence value from combined ones of multiple similarity measurements, wherein the combined ones comprise content, expanded, and category similarity measurements;

providing an objective measure of relevance between the bid term(s) and the site contents as indicated by the confidence value;

analyzing the confidence value to identify the bid term(s); and

using the bid term(s) identified to increase traffic to a site to obtain site exposure;

wherein the computer-executable instructions for generating the category similarity measurements further comprise instructions for:

extracting features from Web site content associated with the directory data, the features comprising a combination at least one of title, metadata, body, hypertext link(s), visual feature(s), and summarization by page layout analysis information; reducing dimensionality of the features via feature selection; categorizing the features via a classifier model to generate the similarity classifier;

generating respective term vectors from the bid term(s), the site contents, and the expanded terms; and

calculating similarity between the respective term vectors as a function of the similarity classifier to determine the category similarity measurements.

Claim 25

A computing device for verifying relevance between terms and Web site contents, the computing device comprising:

a processor; and

a memory coupled to the processor, the memory comprising computer-program instructions executable by the processor for:

retrieving site contents from a bid URL;

formulating expanded term(s) comprising at least one of semantically or contextually related to bid term(s),

generating content similarity and expanded similarity measurements from respective combinations of the bid term(s), the site contents, and the expanded terms, wherein the similarity measurements indicate relatedness between respective ones of the bid term(s), site contents; or expanded terms;

calculating a confidence value from combined ones of multiple similarity measurements, wherein the combined ones comprise content, expanded, and category similarity measurements;

providing an objective measure of relevance between the bid term(s) and the site contents as indicated by the confidence value;

analyzing the confidence value to identify the bid term(s); and

using the bid term(s) identified to increase traffic to a site to obtain site exposure;

wherein the computer-executable instructions for generating the category similarity measurements further comprise instructions for:

extracting features from web site content associated with the directory data, the features comprising a combination of at least one of title, metadata, body, hypertext link(s), visual feature(s), and summarization by page layout analysis information;

reducing dimensionality of the features via feature selection;

categorizing the features via a classifier model to generate the similarity classifier;

generating respective term vectors from the bid term(s), the site contents, and the expanded terms; and

calculating similarity between the respective term vectors as a function of the similarity classifier to determine the category similarity measurements.

Claim 36

A computing device for verifying relevance between terms and Web site contents, the computing device comprising:

retrieving means to obtain site contents from a bid URL;

formulating means to identify expanded term(s) comprising at least one of semantically or contextually related to bid term(s),

generating means to create content similarity and expanded similarity measurements from respective combinations of the bid term(s), the site contents, and the expanded terms, wherein the similarity measurements indicate relatedness between respective ones of the bid term(s), site contents, or expanded terms;

calculating means to determine category similarity measurements between the expanded terms and the site contents in view of a similarity classifier, wherein the similarity classifier has been trained from mined web site content associated with directory data;

calculating means to generate a confidence value from combined ones of multiple similarity measurements, wherein the combined ones comprise content, expanded, and category similarity measurements, wherein the confidence value provides an objective measure of relevance between the bid term(s) and the site contents;

analyzing means to analyze the confidence value to identify the bid term(s); and

increasing means to increase traffic to a site by using the bid term(s) identified;

wherein the generating means further comprise:



extracting means to obtain features from Web site content associated with the directory data, the features comprising a combination of at least one of title, metadata, body, hypertext link(s), visual feature(s), and summarization by page layout analysis information;

reducing means to lessen dimensionality of the features via feature selection;

categorizing means to organize the features via a classifier model to generate the similarity classifier;

generating means to create respective term vectors from the bid term(s), the site contents, and the expanded terms; and

calculating means to identify similarity between the respective term vectors as a function of the similarity classifier to determine the category similarity measurements.

Claim 14, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 15, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 16, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 17, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 19, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 20, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 21, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 22, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 23, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

Claim 24, line 1, change "A computer-readable medium as recited in" to - - A computer-readable storage medium as recited in - -.

### **REASONS FOR ALLOWANCE**

With respect to claims 1-5, 7-17, 19-29, 31-39 and 41-45 are allowed over the prior art of record.

The following is an examiner's statement of reasons for allowance:

With respect to claims 1-5, 7-17, 19-29, 31-39 and 41-45, the claimed features "generating content similarity and expanded similarity measurements from respective combinations of the bid term(s), the site contents, and the expanded terms, wherein the similarity measurements indicate relatedness between respective ones of the bid term(s), site contents, or expanded terms; calculating means to generate a confidence value from combined ones of multiple similarity measurements, wherein the combined ones comprise content, expanded, and category similarity measurements, wherein the confidence value provides an objective measure of relevance between the bid term(s) and the site contents; reducing dimensionality of the features via feature selection; calculating similarity between the respective term vectors as a function of the similarity classifier to determine the category similarity measurements" in conjunction with other elements of the independent claims would not have been found to be anticipated or obvious over the prior art made of record.

The dependent claims, being definite, further limiting, and fully enabled by the specification are also allowed.

The closest prior art, Wyard et al., U.S. Patent No. 6,167,398 relates to an internet information agent which analyses candidate documents for dissimilarity with a reference corpus identified by a user of the agent. Thorsten Joachims, relates to optimizing search engines using clickthrough data but fail to teach the above limitations.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

### CONTACT INFORMATION

2. Any inquiry concerning this communication or earlier communications from the examiner should be directed to JEAN B. FLEURANTIN whose telephone number is 571 – 272-4035. The examiner can normally be reached on 7:05 to 4:35.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, JOHN E BREENE can be reached on 571 – 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 703-308-6606.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).



Jean Bolte Fleurantin

Patent Examiner

Technology Center 2100

April 25, 2007